DOCUMENT RESUME

ED 066 132 HE 003 299

AUTHOR	Pace, C. Robert
TITLE	Thoughts on Evaluation in Higher Education.
INSTITUTION	American Coll. Testing Program, Iowa City, Iowa.
PUB DATE	Feb 72
NOTE	19p.; Speech presented at the invitation of the American College Testing Program and the College of Education, The University of Iowa, Iowa City, April 26, 1971
AVAILABLE FROM	ACT Publications, P.O. Box 168, Iowa City, Iowa 52240
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Administrator Role; *Educational Administration; Educational Change; *Educational Development; Educational Planning; *Evaluation; *Higher Education; Program Evaluation

ABSTRACT

Educational evaluation has in the past been primarily associated with measurement, achievement testing, pupil progress, instructional methods and curriculum. Educational testing and educational research have provided the instruments and models for the evaluation of educational programs. More recently, however, there has been an increased concern about the effectiveness of large scale social action enterprises. A concept of evaluation appropriate for the study of large and complex institutions can be summarized briefly as follows: (1) it begins with the central question "What are the consequences?" rather than the limiting question "What are the objectives?" (2) its style of inquiry is more aptly characterized by the work exploration than by words such as "control" and "focus"; (3) it sees the role of the evaluator as that of social scientist rather than teacher, missionary, reformer, or staff officer to the practitioners; and (4) its purpose is to provide more complex bases for informed judgment. (HS)

ED 066132

ESSAYS ON EDUCATION

NUMBER ONE IN A SERIES:

Thoughts on Evaluation in Higher Education

C. ROBERT PACE

University of California, Los Angeles

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

HE 003 299

THE AMERICAN COLLEGE TESTING PROGRAM

THOUGHTS ON EVALUATION IN HIGHER EDUCATION*

C. Robert Pace
University of California, Los Angeles

*A talk given April 26, 1971, in Iowa City, Iowa, at the invitation of The American College Testing Program and the College of Education, The University of Iowa.

The American College Testing Program is dedicated to the enrichment of education. It was founded as an inviolate public trust and operates as a nonprofit corporation governed by educational representatives from individual states or regions and a Board of Trustees.

A fundamental goal of The Program is to exercise educational leadership by conducting testing, information gathering, evaluating, and related activities in order to (1) assist in the identification and solution of educational problems and (2) communicate to the general and professional publics knowledge and ideas about education.

The chief beneficiaries of The Program's services are students, secondary schools, institutions of higher education, and educational researchers.

Published February 1972
by The American College Testing Program

For additional copies write:
ACT Publications, P.O. Box 168
Iowa City, Iowa 52240

THOUGHTS ON EVALUATION IN HIGHER EDUCATION

The title of this talk is a reminder that national and international events often have repercussions on the campus. Owing to a Ping-Pong game, the People's Republic of China and the Thoughts of Chairman Mao have come into our consciousness. I was in the midst of preparing a well-structured coherent speech when I realized that a new style of communication may become common, the style of Mao's Thoughts, and that perhaps I could contribute a little bit to our readiness for this mode by trying to emulate it. My thoughts, like those of the Chinese, may appear disconnected, underdeveloped, ambiguous, possibly contradictory, and more or less inscrutable. But of course, like the Chinese thoughts, there just may be some underlying significance to them.

Thoughts on the New Direction for Evaluation

In education, during the past 30 years or so, evaluation has been primarily associated with measurement, achievement testing, pupil progress, instructional methods and curriculum. Educational testing and educational research have provided the instruments and the models for the evaluation of educational programs. Some of the most distinguished of these instruments and models have been developed by men associated with The University of Iowa—E. F. Lindquist, Robert Ebel, and William Coffman.

More recently, however, there has been an increased concern about the effectiveness of large scale social action enterprises. So now we see efforts to evaluate a school system, a nationwide head start program, hospitals, welfare, mental health programs, transportation systems, higher education in the United States, and other complex phenomena.

If we take on these larger problems, then clearly the new direction for evaluation in education is toward finding ways of dealing with matters that are not just educational, that have broad social consequences, and that are characterized by complex interactions and often by conflicting objectives and values.

Thoughts on a Concept of Evaluation for Dealing with Complex Phenomena

If educational evaluators are to deal effectively with large problems—such as the effectiveness of a total institution, of a class of institutions, or of higher education in the United States—then they need a concept of evaluation that is

both more comprehensive and more flexible than the familiar experimental model. For this level of complexity and reality, the evaluator must ask different questions, proceed in a different style, and have a new view of his role and purpose.

The central question is not "What are the objectives?" The central question is "What are the consequences?" If one should ask himself the question "What are the objectives of the United States in Vietnam?" no doubt some answers would come to mind. But if instead one asked himself the question "What are the consequences of the war in Vietnam?" a much greater range of inquiry immediately would be suggested and required. It seems obvious that the range of one's inquiry is guided by the questions one asks. "Objectives" are a subheading under "consequences." Of course, one hopes that among the consequences are some which are intended as objectives, but looking at the extent to which objectives are achieved will not answer the larger question about consequences. Therefore, the first requirement for a new model of evaluation is to begin with the question "What are the consequences?"

The second necessary element in a new model is one which relates to the appropriate style of inquiry. An apt term for this is "exploration." Traditionally, the style of inquiry has been characterized by the words "control" and "focus." Exploration is a freer style—one which encourages hunches, is uncommitted, and seeks discovery. If the program one hopes to evaluate is continually changing in methods, materials, personnel, and subjects, this does not mean that it cannot be evaluated. On the contrary, one may discover that programs that are being modified continually are more effective than programs that remain relatively static. The spirit of the evaluator should be adventurous. If only that which could be controlled or focused were evaluated, then a great many important educational and social developments would never be evaluated—at least not by "evaluators"; and that would be a pity. To suggest that the style of the controlled experiment needs to be replaced by an exploratory style does not mean that one's approach should be any less careful or rigorous. Exploration involves searching, probing, and testing alternatives and interactions. It can be tough-minded and theory-based. But the word *exploration* also connotes a freedom to look around, to seek new measures and methods, and to value ingenuity and curiosity.

The third element for a revised concept of evaluation is a new view of the role of the evaluator and the purpose of evaluation. Historically, many people have seen evaluation as an instrument of reform. The reason for evaluating any present activity or program was to improve it. As a result of this view,

the parties to the activity or program had to be involved in its evaluation because their very involvement would increase the likelihood that they would be willing to change in the light of the findings. Thus, the process of carrying out an evaluation—group participation and cooperation—was directly related to achieving the purpose of evaluation, namely, change and improvement. Implicitly, the role of the evaluator was in some respects the role of missionary and reformer. Another common view is that the purpose of evaluation is to give feedback to the decision makers (presumably teachers and administrators). In this case the role of the evaluator is that of a kind of staff officer to the practitioner. Neither of these views seems quite appropriate for evaluations that deal with large, complex social or educational programs such as higher education in the United States. A more appropriate concept of evaluation would define the role of the evaluator as that of a social scientist and the purpose of evaluation as that of providing more complex bases for informed judgment. When asked about the relationship between social scientists and administrators, Harold Lasswell is reported to have said that the role of the social scientist is to complicate the tasks of the decision makers. The evaluator, as a social scientist, should play this role.

A concept of evaluation appropriate for the study of large and complex institutions can be summarized briefly as follows:

1. It begins with the central question "What are the consequences?" rather than with the more limiting question "What are the objectives?"

2. Its style of inquiry is more aptly characterized by the word "exploration" than by the words "control" and "focus."

3. It sees the role of the evaluator as that of a social scientist rather than that of a teacher, missionary, reformer, or staff officer to the practitioners.

4. Its purpose is to provide more complex bases for informed judgment.

In passing, I mentioned the words "decision" and "decision maker." There is a strong move today to relate evaluation to decision making; to argue indeed that this is the only real purpose of evaluation. I believe this may be a wrong direction to take, at least in the sense that there are hazards in such an emphasis. So, I turn next to:

Thoughts on Evaluation and Decision Making

The emphasis on "decision maker" has, I think, several potential dangers.

1. It suggests doing something practical and specific for one person or group of persons.

2. It suggests a close relationship between evaluator and administrator. I personally think it is advisable to maintain a firm independence of the two—and that evaluation should not be defined in a way that makes independence difficult to attain.

3. It sets up an expectation that the proof of a good evaluation is whether the decision maker acts in accord with the evaluator's results. One could get a reputation for being a good evaluator by being able to persuade most of his clients to act on his advice; and this has no necessary connection with the content or quality of the evaluation.

I see no objection to the idea of there being some ultimate decision points *in the mind of* the evaluator as he carries out his work. This may be a useful distinction between evaluation and some kinds of research. If one has a realization that the purpose of what he does is to provide the best possible basis for informed judgments (decisions?), his thinking about his task will surely be influenced. And this will be a different influence from that which operates on the researcher whose purpose is to discover or explain some phenomenon.

Consider the subtle and not so subtle influence on one's thinking that might stem from each of the following key orientations: decision, explanation, judgment.

Decision suggests determination, deciding between alternatives, making up one's mind. Explanation suggests to account for, to make plain, to make clear the cause or reason, and to interpret. Judgment suggests good sense, discretion, to discern circumstances and draw conclusions, to judge probable consequences. If one's purpose is decision, one enters the field of science; and if judgment is the purpose, one enters the field of values. Decision is the realm of administration; explanation is the realm of science; judgment is the realm of evaluation.

In my view, "decision making" is too narrow a focus for describing the purpose and role of evaluation; explanation is too abstract, impersonal, and,

in the history of educational research, has been too limited in what it has explained; and so I am left with "judgment" as the central orientation for evaluation studies, based on the proposition that wise judgment demands an awareness of complexity and consequences, a consideration of values, and the possession of information relevant to such complexities, consequences, and values.

I turn next to some propositions about the nature of systems.

Thoughts on the Nature of Complex Systems

1. All schools or systems or large programs consist of many different activities.

   At the individual level, a college student hears lectures, participates in discussions, takes tests, reads books, writes papers, talks with professors, participates in or attends sporting events and cultural events, meets people, makes new friends, joins organizations, falls in love, etc.

   At the institutional level there are curricula, courses, academic and nonacademic personnel with various duties, research activities, public service activities, fund raising activities, political or lobbying activities, building and planning activities, administrative routines, etc.

   This proposition is so obviously true in the observations and experience of everyone that no further documentation is needed.

   The implication of it is simply that an evaluation must take account of this range of activities. Data reduction is often useful. But one needs to realize that data enrichment is also essential.

2. **All programs have multiple consequences, many of which are not objectives or intentions of the programs.**

   The new high school physics has produced students better prepared for college physics; but there are proportionately fewer high school students who study physics today than there were formerly.

   School grades are meant to be a reward for achievement, but the more importance that is attached to grades, the more frequently cheating behavior is likely to occur.

Any learning task may have affective results that are not compatible with a desire for any further learning of that task.

The conclusion to be drawn from this proposition is that an evaluation which considers only the extent to which program objectives are attained will always be an inadequate evaluation and may often be a false or misleading evaluation.

We need suggestions or rules for determining what range of consequences ought to be included in any evaluation study.

3. **All programs or systems have functions that are not the same as program objectives or as program consequences.**

One function that universities perform is to facilitate selective mating, simply by being a place where large numbers of reasonably bright young people have opportunities to meet one another. This is not an objective of higher education, nor a "side effect" of any particular program. But it is probably a beneficial function; and if universities were dispersed, or opportunities for a broad range of acquaintance reduced, the resulting societal detriment might be great.

Another function that universities perform is to keep large numbers of young people out of the labor market. Again, if this function was not served, the societal consequences would be great. But this is not an objective of higher education.

So, in any broad evaluation, one needs to find ways of identifying some of the significant functions that institutions or programs or systems have. And these are quite different from objectives or from the "unintended consequences" of any given program or system.

4. **Single instructional programs, within a school or school system, always interact with other programs in the system.**

A math program, or an art program, or a civics program, or an athletic program is only one part of the total curriculum or set of activities in the system.

Given some finite time period and finite resources, the more time and money devoted to one program the less time and money there will be for another program.

The conclusion to be drawn from this proposition is that the "success" of any program must always be judged in relation to the "success" or "lack of success" of other programs within the system.

5. **All programs exist in a context that is larger than the specific program, or than the sum of programs in the system.**

   Any program or school is part of a larger environment that may have a bearing on the success of the school or program—i.e., the neighborhood or community, different social or ethnic or religious compositions, governmental or other external restraints or benefits, etc.

   Also, programs or schools are inhabited by different sorts of people who may have different attitudes, resources, talents, and orientations—personality, motivation, family backgrounds, interests, abilities, financial conditions, staff training and competence, etc.

   These characteristics of inhabitants and environments may have a bearing on what programs within the system are successful and on what methods are productive. In an evaluation one needs to identify these possible sources of influence.

6. **When a specific program or product or procedure is introduced into the system the criteria for evaluating it become "system level" criteria—that is, the criteria used initially in evaluating the specific product are no longer sufficient.**

   Individually Prescribed Instruction (IPI) or Computer Assisted Instruction (CAI) are efficient learning strategies. Introduced as a major part of a school's total program of activity, they may have other consequences—such as retarding socialization, inhibiting creativity, etc.

   Gasoline is tested by octane rating. The higher the rating the better the gas. But this "good" product, when introduced on a large scale into a system of automotive transportation, now needs to be evaluated by different criteria and may then be judged differently—i.e., the higher the octane rating, the greater the air pollution.

   We need to make clear the differences between relevant criteria for different levels of complexity.

7. Evaluation requires placing values on different criteria (objectives, consequences, and functions).

   One cannot duck these value judgments by pleading impartiality or neutrality.

   Even empirically derived "weights" are dependent in the first place on the criterion one uses in deriving them.

   So, what advice can one give to evaluators about the value judgments they make? Is it sufficient to suggest that a broad range of values must be employed and compared? Or that opposing values should always be considered?

   We'll come back to this value issue a little later.

Up to this point I've expressed some thoughts about evaluation of complex phenomena, the role of the evaluator, and the nature of systems. I do not mean to suggest that these thoughts are applicable to all sorts of evaluation problems. Rather, there appears to be a natural congruence or relationship between the nature of the problem and the nature of evaluation that is applicable to it.

Some perspective on different views of evaluation may be gained by attempting to classify the variety of evaluation activities one finds today in a way that acknowledges the validity of each and the validity of the differences between them. These differences are highlighted by presenting a series of contrasting conditions.

Thoughts on Contrasting Concepts of Evaluation

The most important contrast, in the sense that its ramifications are extensive and obvious, is one that relates to the size, complexity, and duration of what is to be evaluated. Consider the following:

When the unit to be evaluated is a small unit—small in size, limited in scope, and short in time—such as a half-hour film, a specific unit of instruction in a single course, a particular method of teaching, or a programmed text,

Then, the following conditions are usually true:

1. The treatment (unit to be evaluated) can be clearly and explicitly defined;

2. The treatment can be compared with alternative treatments or control groups;
3. The requirements of experimental design involving random assignments of subjects to treatments can usually be met;
4. The assumptions for statistical tests of significance, appropriate in a hypothesis testing experiment, can usually be met.

Under these conditions, relevant evaluations can be:

1. Directly related to behaviorally defined objectives;
2. Designed as hypothesis testing experiments;
3. Largely limited to the intended effects of the program or treatment.

In contrast:

When the unit to be evaluated is large, complex, and of long duration—such as a school system, a total institutional program, or higher education in the U.S.—

Then, the following conditions are usually true:

1. The treatment (unit to be evaluated) cannot be clearly and explicitly defined because it is not in fact a unitary phenomenon but is, instead, made up of many units interacting with one another in varied ways and having varied purposes;
2. Gross differences between treatments can sometimes be found and compared, but control groups in the usual experimental sense do not exist;
3. Random assignment of subjects to treatments is impossible except occasionally in some small segment or limited part of the larger treatment;
4. Treatments are constantly undergoing change.

Under these conditions, relevant evaluation:

1. Must consider a broad range of educational and social consequences;
2. Should never be limited by or confined to the stated objectives or intended effects of the program or treatment;
3. Should look for but may not always find contrasting conditions in natural settings for comparative analysis;
4. Must employ different methods of treating data—descriptive, exploratory, multivariate, etc.

Also, as the unit or program to be evaluated becomes larger, the contexts within which the program operates—contexts such as organizational and administrative conditions, the relation to other programs within the school or system, the nature of the clientele and the community, the financial resources and their allocation, the atmosphere of the school—have a greater opportunity for influence; and it becomes crucial

to include a range of such potentially relevant contextual variables in one's evaluation design.

As size, scope and duration change—from small to large, simple to complex, short to long—there are corresponding changes in the nature and the procedures of evaluation, relevant to the differing conditions.

One can also classify and suggest the implications of different concepts of the role of the evaluator and the purpose of evaluation. Again, consider the following contrasting cases.

When the evaluator is basically a teacher, reformer, or staff officer to the practitioner and the purpose of evaluation is to improve or change a program or practice,

Then, the process of evaluation is characterized by:

1. A client-centered orientation—in that the clients specify the objectives (usually with help from the evaluators);
2. A cooperative mode of inquiry—in that the clients or practitioners, in addition to the evaluators, plan, conduct, and interpret the inquiry.

The intended result is decision and action.

But when the evaluator is seen as a neutral social scientist and the purpose of evaluation is information and analysis,

Then, the process of evaluation is characterized by:

1. An independent orientation—in that the range of inquiry includes but is not limited to the client's intended objectives;
2. A collaborative mode of inquiry—in that expertise from relevant disciplines is brought to bear on the design, conduct, and analysis of the inquiry.

The intended result is the provision of more complex bases for informed judgment.

A similar classification or contrast can be considered in relation to the nature and place of "decision" in an overall conceptualization of evaluation. To indicate the ends of a continuum, one might say that:

1. *When* one is concerned with a specific limited topic that is generally within the control and responsibility of a single individual or small group, and that can be stated as a choice between clearly defined alternatives—*then,* it may be reasonable and useful to view evaluation as directly contributory to the action of a decision maker.

2. *But when* one is concerned with a complex topic—interrelated with other topics, of relevance and importance to a variety of individuals and groups, and involving several plausible and perhaps overlapping alternatives—*then*, it seems more appropriate and useful to view evaluation as generally contributory to informed judgment and policy direction.

We've been talking about how evaluation practice is necessarily related to what is being evaluated. Now we come to some possibly more important and interesting thoughts.

Thoughts on How One's Practice of Evaluation Reflects, Consciously or Not, One's Concept of Education

There can be little doubt that higher education is being judged widely and vigorously by the public, by students, and by the profession itself. Put another way, like it or not, higher education is being evaluated. It is therefore critical to understand how well it is being evaluated: Who is making what observations? What measures are applied? What goals or purposes are assumed? More importantly, what concepts about the enterprise of education itself are determining what to look for and how to interpret the results? If education is changing—reassessing old values, exploring new directions, serving new purposes—then the scope of evaluation must be broad enough to include both the new and the old. Evaluation is both the process and the result of judging the effects of an educational program. If the process is limited (restricted to certain types of observation) the results will be misjudgment. In times of social change there is a particularly strong need for broad perspective and wise judgment. If the tools and concepts of evaluation practice are not now adequate for this responsible and difficult task, then we must devise better tools and more relevant concepts.

Thoughts on Changes in Higher Education

The major strains and stresses of higher education today are largely the product of very rapid and profound events within the past 25 years. Twenty-five years ago there was almost no large scale government funded research on university campuses; there was almost no federal support for student financial assistance nor for campus construction; junior colleges were few and generally not well accepted; minority groups were conspicuously absent from college campuses; there were more students enrolled in private colleges than in public ones; and altogether only about 1 out of 9 young

people of college age were in college. Today 5 out of 9 young people of college age are in colleges; and about 60% of all high school graduates in the U.S. go to colleges; nearly three-fourths of all college students are in public institutions, not private ones; junior colleges abound; minority groups are demanding admittance to and being encouraged to attend college; programs for financial aid to students (the G.I. Bills, traineeships and fellowships in various fields, and low interest loans) are common; college construction has been facilitated; and research is a prominent activity in the leading universities. The task of absorbing and adapting to these changes has been enormous.

Ordinarily one would advise caution in thinking about the next 25 years; but to counsel caution today may be more hazardous than to advocate risk. It is possible that, within the next 10 years or so, we shall see college programs that are as much concerned with personal development, human relations, problem solving, and social betterment as with conveying the knowledge and methods of particular academic disciplines; that we shall have made a national political commitment to universal higher education; and that as a consequence of these changes we shall see, within the academic community, a more broadly relevant set of priorities and programs that will bring the universities and the larger society into closer and more productive relationship in a technologically dependent but humanely guided world.

The more "humanely guided world" is the key to other major value changes—such as a reemphasis on quality (quality of the environment, quality of life, quality of personal relations, quality of work, etc.) and some emergent belief that one cannot really have both quantity and quality (at least not in some areas). There is a certain anti-science, anti-technology attitude that is quite pervasive. This in turn is related to attitudes of anti-bureaucracy, impersonality, and efficiency—the encroachment of the super administrative State upon the lives and individuality and privacy and values of everyone. Dehumanization. The final dominance of the Pentagon of Power, as Lewis Mumford would put it.

It is the dominance of the mechanical and quantitative sciences that is being questioned and attacked. This questioning comes from some of our most respected scholars. Here, for example, are four quotes from Lewis Mumford's recent book. They refer to the emergence of mechanistic views of the universe. And although in the physical sciences today the most advanced researchers and theorists are less mechanistic, the myth of the machine still dominates popular culture.

Galileo committed the crime of giving up man's birthright—man's remembered experience, or accumulated culture. He had no notion that his radical distinction between the external world and the internal world, between the objective and the subjective, between the quantitative and the qualitative, between the mathematically describable, and thus knowable, and the irreducible, inaccessible, unanalyzable, and unmeasurable, was a false distinction, once human experience, in its symbolized fullness—itself a deposit of countless ages of organic life—was left out of account.

If the new science had begun with the observer himself as an essential component in his own scheme, the insufficiency of his mechanical model and his denatured and dehumanized universe would have been apparent—indeed, inescapable. Without intuitions and memories, without ancient cultural landmarks, the intelligence is enfeebled, and the report it gives on its own say so is so incomplete, so qualitatively inadequate, so structurally distorted that it becomes downright false.

What made the new world picture so potent was that its method of deliberately ignoring the complex reality of organisms was an immense labor saving device: its pragmatic efficiency counterbalanced its conceptual superficiality.

All these achievements made the mechanical world picture highly acceptable. In every department, the sign of quantity or magnitude would, ideally, become a necessary part of every qualitative judgment. Up to a point, then, the new method was self-validating. It was only when it concentrated on quantity to the exclusion of form, pattern, functional organization, design, that the weakness of the emphasis upon so-called primary qualities became a handicap. Those who developed the mechanical world picture ignored Liebnitz's salient distinction between *accurate knowledge* and *adequate knowledge* . . . .

The responsibility for *adequate knowledge* is one that evaluators must accept.

Thoughts on Analogies and Models of Education and Evaluation

Many of the analogies and models we use in thinking about education and evaluation are drawn from fields that have no necessary connection with the nature and quality of education. Higher education is not a factory, receiving raw material, processing it, and turning out products having certain performance characteristics. Nor is higher education a business, distributing goods at so much per unit cost. Nor is it a bureaucracy run by bosses, with flow charts, communication networks, decision points, and job descriptions. But these mechanical and administrative analogies have their counterparts in the language of educational research and evaluation—the measurement of input-output differences, specified performance objectives, college effects, test score gains, etc. Most recently the popular terminology includes behavioral objectives, product development, cost effectiveness, performance contract, management information systems, and accountability. There are potential dangers in these conceptualizations, methods, and measures: for they tend to emphasize and reinforce an administrative and efficiency view of

the nature and purpose of education. Evaluation, if it follows administrative, management, and efficiency models, can and probably will contribute to the industrialization of higher education—education in which order, duty, responsibility, efficiency, and performance are the dominant values.

Surely one of the objects of student disaffection with higher education is precisely this matter—impersonality, bureaucracy, dehumanization—and indeed it is a central object of their protest against the larger society. To reverse what is judged to be a distorted inversion of values we need to develop more humane and organic analogies and models.

The relevant analogies are biological, ecological, organic, psychological, sociological, and philosophical. A college or university is a habitat, a society, a community, an environment, an ecosystem. It should be judged by the quality of life that it fosters, the opportunities for experience and exploration it provides, the concern for growth, for enrichment, and for culture that it exemplifies. The question is not just "what does your machine produce?" but also "how does your garden grow?"

College could be conceptualized as an environment for exploration—of self, of knowledge and skills, of ideas and values, of society, conscience, community, and commitment. The quality of the exploration and the character of the environment are inseparable. But we have no theory or methodology of evaluation that is compatible with this emergent value system in education or in society.

Thoughts on Some Unmeasured Qualities of Experience and Environment

What are the words one would use to describe the quality of an experience or the character of an environment?

breadth-depth
pervasive-limited
intense-bland
expansive-restrictive
structured-unstructured
permissive-demanding
personal-impersonal
aggressive-cooperative
friendly-hostile

These are not the words one commonly uses to describe achievement or production. They suggest new challenges for evaluation.

Consider some recent conclusions based, I think, on inadequate observations and on limited concepts of education.

It is said, for example, that Ivy League, selective institutions have less impact or effect than colleges of lesser prestige. "Good" colleges are not as effective as "poor" colleges! Such conclusions are based on achievement tests and personality tests. They are, in part, an artifact of the measuring instruments; but more profoundly, they reflect an "achievement-gain-progress" concept of education to the exclusion of other concepts.

Is there not also a unique, once-in-a-lifetime experience about going to college, an experience whose value is not assessed by present tests? College has a tremendous impact on many people who experience it. There is a flood of testimony to this effect—and especially in the so-called elite institutions. If our more formal measures don't show it, then perhaps we've just not looked for the right indicators.

Thoughts on Evaluation as the Conscience of Education

Martin Meyerson once said that the role of an evaluation center was to make the higher education community educationally self-conscious.

Lester Anderson has amplified this by stating that through research and scholarship it strives to awaken the consciousness to opportunities, enhance sensitivity to the consequences of decisions, suggest alternatives, and foster introspection and debate on the part of faculty members and administrators.

With these thoughts I heartily agree.

Beyond the contributions that evaluation has made in the service of management and economy, and educational efficiency and productivity, there is another contribution that has not yet been fully explored or pursued—evaluation in the service of conscience, promoting a heightened awareness of consequences, a more critical awareness of values, and a deeper concern for the qualities of experience.

ABOUT THE AUTHOR

Dr. C. Robert Pace is professor of higher education at the University of California at Los Angeles (UCLA) and director of the higher education evaluation program in the Center for the Study of Evaluation, a U.S. Office of Education (USOE) funded research and development center.

Portions of Dr. Pace's speech, with some modifications, have been drawn from two papers published by the Center for the Study of Evaluation: CSE Report No. 8, "Evaluation Perspectives: '68" (Pace, 1969) and CSE Report No. 51, "A National Evaluation of Higher Education: Plans and Perspectives" (Pace, 1969).

Dr. Pace has served as a consultant to the Research and Development Division of The American College Testing Program (ACT).

The views expressed are those of the writer and do not imply endorsement by USOE or ACT.

ESSAYS ON EDUCATION

NUMBER ONE IN A SERIES: C. ROBERT PACE